

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:22:47 ; Search time 18.74 seconds
(without alignments)
243.889 Million cell updates/sec

Title: US-09-619-198-7

Perfect score: 292

Sequence: 1 QAEATRQAAQAEERLADLAS.....GLQETQQREREREAEQE 60

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	292	100.0	617	2	I56530	gene VGF protein - rat
2	156	53.4	711	2	S05381	VGF8a protein precursor
3	75.5	25.9	568	2	S15008	gene disco protein
4	73.5	25.2	622	2	S71342	calnexin precursor
5	70	24.0	1979	2	C71622	hypothetical protein
6	68	23.3	966	2	S25365	CYC8 protein - yeast
7	68	23.3	1353	2	T27404	hypothetical protein
8	67	22.9	478	2	S56904	hypothetical protein
9	67	22.9	913	2	T52485	neurofilament protein
10	66.5	22.8	993	2	T17230	hypothetical protein
11	66.5	22.8	4687	1	A39638	plectin - rat
12	65.5	22.4	284	1	TWCHA	tropomyosin 2, ske
13	65.5	22.4	284	2	A28499	tropomyosin alpha
14	65.5	22.4	284	2	S24402	alpha-tropomyosin
15	65.5	22.4	284	2	A26113	tropomyosin alpha
16	65.5	22.4	284	2	S24401	tropomyosin 2, ske
17	65	22.3	120	2	F72586	hypothetical protein
18	65	22.3	1023	2	S12519	glutactin - fruit
19	64.5	22.1	1038	2	T02634	rep protein homolog
20	64	21.9	92	2	S23527	embryonic abundant
21	64	21.9	544	2	B44841	low molecular weight
22	64	21.9	707	2	T01803	hypothetical protein
23	63	21.6	504	2	S51942	prunin 2 precursor
24	63	21.6	537	2	A54424	acrosomal protein
25	63	21.6	678	2	F70404	flagellar export protein
26	63	21.6	1018	2	S44758	cl4B9.6 protein - rat
27	62.5	21.4	175	2	S57987	hypothetical protein
28	62.5	21.4	320	2	S50716	hypothetical protein
29	62.5	21.4	462	2	I53414	lamin C2 - mouse

30	62.5	21.4	572	1	VEHULC	lamin C - human
31	62.5	21.4	574	2	S04333	lamin C - mouse
32	62.5	21.4	664	1	VEHULA	lamin A - human
33	62.5	21.4	665	2	S27267	lamin A - rat
34	62.5	21.4	665	2	S28182	lamin A - mouse
35	62.5	21.4	1184	2	G01763	atrophin-1 - human
36	62.5	21.4	1560	2	T30282	calcium-binding protein
37	62	21.2	102	2	T01285	hypothetical protein
38	62	21.2	536	1	A47190	transducer protein
39	62	21.2	536	2	E84318	Htr1 transducer protein
40	62	21.2	924	2	S06117	myosin heavy chain
41	62	21.2	1110	2	I51116	NF-180 - sea lamp
42	62	21.2	1432	2	B85431	trichohyalin like
43	62	21.2	2007	1	B43402	myosin heavy chain
44	61.5	21.1	1087	2	T30330	gelosolin-related protein
45	61.5	21.1	1093	2	A47212	transcription factor

ALIGNMENTS

RESULT 1

I56530

gene VGF protein - rat

N:Alternate names: VGF

C:Species: Rattus norvegicus (Norway rat)

C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 01-Dec-2000

C:Accession: I56530; A39748; JH0454

R:Rawley, R.J.; Scheibe, R.J.; Wagner, J.A.

J. Neurosci. 12, 2573-2581, 1992

A:Title: NGF induces the expression of the VGF gene through a cAMP or response element

A:Reference number: I56530; MUID: 92309005

A:Accession: I56530

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-617 <RES>

A:Cross-references: GB:M74223; NID:q207650; PIDN:AAA42336.1; PID:q207651

R:Salton, S.R.J.; Fischberg, D.J.; Dong, K.W.

Mol. Cell. Biol. 11, 2335-2349, 1991

A:Title: Structure of the gene encoding VGF, a nervous system-specific mRNA that is r

A:Reference number: A39748; MUID: 91203852

A:Accession: A39748

A:Molecule type: DNA

A:Residues: 1-175, 'D', 177-617 <SALI>

A:Cross-references: GB:M60522

R:Salton, S.R.J.

J. Neurochem. 57, 991-996, 1991

A:Title: Nucleotide sequence and regulatory studies of VGF, a nervous system-specific

A:Reference number: JH0454; MUID: 91318308

A:Accession: JH0454

A:Molecule type: mRNA

A:Residues: 1-175, 'D', 177-617 <SALI>

A:Cross-references: GB:M60525

A:Experimental source: cultured cell PC12, clone NGF33.1

C:Comment: This protein is induced in PC12 cells to a greater extent by nerve growth

C:Genetics:

A:Gene: VGF

C:Keywords: growth factor

Query Match 100.0% Score 292; DB 2; Length 617;
Best Local Similarity 100.0%; Pred. No. 3.7e-23;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QAEATRQAAQAEERLADLASDLLLQYLLQGARGDRDLGGRLQETQQREREREAEQE 60

Db 313 QAEATRQAAQAEERLADLASDLLLQYLLQGARGDRDLGGRLQETQQREREREAEQE 372

RESULT 2

S05381

VGF8a protein precursor - rat

C:Species: Rattus norvegicus (Norway rat)

;;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 13-Nov-1998
;Accession: S05381
;Species: Rana rugosa (Korean frog)
;Title: A protein induced by NGF in PC12 cells is stored in secretory vesicles and released by exocytosis
;Reference number: S05381; MUID: 90005425
;Status: not compared with conceptual translation
;Molecule type: mRNA
;Residues: 1-711 <POS>
;1-23/Domain: signal sequence #status predicted <SIG>
;24-711/Product: VGF8a protein #status predicted <MAT>

Query Match 53.4%; Score 156; DB 2; Length 711;
Best Local Similarity 46.9%; Pred. No. 7.4e-09;
Matches 38; Conservative 7; Mismatches 10; Indels 26; Gaps 2;

2y 5 TRQAAQERLADLASLLQYLQ-----GGARQD--LG 38
y 316 TRQAAQERLADLASLLQYLQAAITAISGVAGCRRRSKSGREGGRRRDVAVG 375
39 GRGLQEQERENEEAEQ 59
376 GRGGEIETQASEQERAAER 396

Db

RESULT 3
315008
gene disco protein - fruit fly (Drosophila melanogaster)
;Species: Drosophila melanogaster
;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 21-Jul-2000
;Accession: S15008; S14561
;Heilig, J.S.; Freeman, M.; Laverty, T.; Lee, K.J.; Campos, A.R.; Rubin, G.M.; Steller, R.M.O. 10, 809-815, 1991
;Title: Isolation and characterization of the disconnected gene of Drosophila melanogaster
;Reference number: S15008; MUID: 91184126
;Accession: S15008
;Molecule type: DNA
;Residues: 1-568 <HEI>
;Cross-references: EMBL:X56232; NID:g7849; PIDN:CAA39689.1; PID:g7850
;Gene: disco
;Cross-references: FlyBase:FBgn0000459
;Keywords: DNA binding; zinc finger

Query Match 25.9%; Score 75.5; DB 2; Length 568;
Best Local Similarity 42.3%; Pred. No. 1.6;
Matches 22; Conservative 4; Mismatches 19; Indels 7; Gaps 1;

2y 9 AAQERLADLASLLQYLQGGARQDGLGRLQEQERENEEAEQ 60
Db 292 AADEERAADQLDFSLH-----KRRKSEQDREQEQERENEEAEQ 336

RESULT 4
371342
calnexin precursor - Korean frog
;Species: Rana rugosa (Korean frog)
;Date: 29-Jan-1998 #sequence_revision 06-Feb-1998 #text_change 20-Jun-2000
;Accession: S71342
;Yamamoto, S.; Nakamura, M.
;FEBS Lett. 387, 27-32, 1996
;Title: Calnexin: its molecular cloning and expression in the liver of the frog, Rana rugosa
;Reference number: S71342; MUID: 96234004
;Accession: S71342
;Molecule type: mRNA
;Residues: 1-622 <YAM>
;Cross-references: EMBL:D78590; NID:gl514958; PIDN:BA11426.1; PID:gl514959
;Experimental source: liver
;Superfamily: calnexin
;Keywords: calcium binding; endoplasmic reticulum; molecular chaperone; transmembrane p

F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-622/Product: calnexin #status predicted <MAT>
F;478-500/Domain: transmembrane #status predicted <TM>

Query Match 25.2%; Score 73.5; DB 2; Length 622;
Best Local Similarity 30.0%; Pred. No. 2.8;
Matches 18; Conservative 10; Mismatches 25; Indels 7; Gaps 1;

Qy 1 QAEATROAAQERLADLASLLQYLQGGARQDGLGRLQEQERENEEAEQ 60
Db 546 EEKAKKADQEDNAEQAEK-----QTGEGGAGAGQSGEEEEEEEEEE 598

RESULT 5
C71622
hypothetical protein PFB0145c - malaria parasite (Plasmodium falciparum)
;Species: Plasmodium falciparum
;Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
;Accession: C71622
;R;Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, R. Science 282, 1126-1132, 1998
;Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
;Reference number: A71600; MUID: 99021743
;Accession: C71622
;Status: preliminary; nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-1979 <GAR>
;Cross-references: GB:AE001375; GB:AE001362; NID:g3845105; PIDN:AAC71819.1; PID:g384
;Experimental source: clone 3D7
;Genetics:
;Gene: PFB0145c

Query Match 24.0%; Score 70; DB 2; Length 1979;
Best Local Similarity 44.7%; Pred. No. 21;
Matches 17; Conservative 9; Mismatches 10; Indels 2; Gaps 1;

Qy 23 LLOYLGGARQDGLGRLQEQERENEEAEQ 60
Db 314 LLEYEEKKNQIDLLNK--QEKEKEKEKEKEKE 349

RESULT 6
S25365
CYC8 protein - yeast (Saccharomyces cerevisiae)
;Alternate names: glucose repression mediator; protein YBR0908; protein YBR112c; SSN
;Species: Saccharomyces cerevisiae
;Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 11-Jan-2000
;Accession: S25365; S48277; S45980; S25404; S25405; A30906; S44692
;Mannhaupt, G.; Stucka, R.; Ehnl, S.; Vetter, I.; Feldmann, H.
;Yeast 8, 397-408, 1992
;Title: Molecular analysis of yeast chromosome II between CMD1 and LYS2: the excision
;Reference number: S25364; MUID: 92327848
;Accession: S25365
;Molecule type: DNA
;Residues: 1-966 <MAN>
;Cross-references: EMBL:X56247; NID:g3548; PIDN:CAA46973.1; PID:g3550
;Mannhaupt, G.; Stucka, R.; Ehnl, S.; Vetter, I.; Feldmann, H.
;Yeast 10, 1363-1381, 1994
;Title: Analysis of a 70 kb region on the right arm of yeast chromosome II.
;Reference number: S48255; MUID: 95208357
;Accession: S48277
;Status: nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-966 <MAW>
;Cross-references: EMBL:X78993; NID:g476045; PIDN:CAA55615.1; PID:g476068
;Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1994
;Feldmann, H.; Mannhaupt, G.; Schwarzlose, C.; Vetter, I.
;submitted to the Protein Sequence Database, August 1994
;Reference number: S45927
;Accession: S45980

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:26:07 ; Search time 24.67 Seconds
(without alignments)
355.749 Million cell updates/sec

Title: US-09-619-198-7
Perfect score: 292
Sequence: 1 QAEATROAAQEEERLADLAS.....GLQETQOERENREERAEQ 60
Scoring table: BLOSUM62
Gap 10.0 , Gapext 0.5

Se-ched: 473505 segs, 146272329 residues
Tc number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:
1: sp.archaea:
2: sp.bacteria:
3: sp.fungi:
4: sp.human:
5: sp.invertebrate:
6: sp.mammal:
7: sp.mhc:
8: sp.organelle:
9: sp.phage:
10: sp.plant:
11: sp.rodent:
12: sp.virus:
13: sp.vertebrate:
14: sp.unclassified:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	249.5	85.4	615	4	Q9UDW8	Q9udw8 homo sapien
2	249.5	85.4	616	4	O15240	O15240 homo sapien
3	87.5	30.0	800	5	Q9Y102	Q9y102 drosophila
4	75	25.7	287	12	Q9DWH5	Q9dwh5 rat cytoleg
5	75	25.7	320	11	Q9CKX9	Q9ckx9 mus musculu
6	73.5	25.2	622	13	Q98985	Q98985 rana rugosa
7	72	24.7	897	13	Q13098	Q13098 xenopus lae
8	71	24.3	316	11	O62254	O62254 mus musculu
9	71	24.3	540	11	O62253	O62253 mus musculu
10	70	24.0	238	5	Q9VK98	Q9vk98 drosophila
11	70	24.0	1979	5	O96133	O96133 plasmodium
12	69.5	23.8	1326	2	Q9L2C3	Q9l2c3 streptomyce
13	69	23.6	808	5	O18409	O18409 drosophila
14	68.5	23.5	773	13	P79805	P79805 morone saxa
15	68.5	23.5	3469	5	Q9U4I2	Q9u4i2 drosophila
16	68.5	23.5	3502	5	Q9V1J9	Q9v1j9 drosophila
17	68.5	23.5	3616	13	Q9W6V0	Q9w6v0 gallus gall
18	68	23.3	1353	5	Q9XW65	Q9xw65 caenorhabdi
19	68	23.3	1985	11	Q9JIS7	Q9jis7 mus musculu

20	68	23.3	2037	5	Q9VWH5	Q9vwh5 drosophila
21	67	22.9	316	11	Q9DLE0	Q9dle0 mus musculu
22	67	22.9	316	11	O9CWI3	O9cwi3 mus musculu
23	67	22.9	913	13	O13099	O13099 xenopus lae
24	66.5	22.8	595	4	Q9UE04	Q9ueu4 homo sapien
25	66.5	22.8	695	5	Q9BHM3	Q9bhm3 paramecium
26	66.5	22.8	993	4	Q9UFW3	Q9ufw3 homo sapien
27	66.5	22.8	1504	4	Q9UES6	Q9ues6 homo sapien
28	66.5	22.8	1528	4	Q9Y2I1	Q9y2i1 homo sapien
29	66	22.6	543	4	Q9BY87	Q9by87 homo sapien
30	65.5	22.4	242	10	Q9XE71	Q9xe71 sorghum bic
31	65.5	22.4	283	13	Q91968	Q91968 gallus gall
32	65.5	22.4	283	13	Q90740	Q90740 gallus gall
33	65.5	22.4	517	5	Q9W4E8	Q9w4e8 drosophila
34	65.5	22.4	614	5	P91956	P91956 lytechinus
35	65.5	22.4	1170	4	Q9H0L6	Q9h0l6 homo sapien
36	65	22.3	120	1	Q9YCU9	Q9ycu9 aeropyrum p
37	65	22.3	1026	5	Q9VLJ3	Q9vlj3 drosophila
38	65	22.3	1532	10	Q9L0W1	Q9l0w1 arabidopsis
39	65	22.3	2656	5	Q9NE92	Q9ne92 leishmania
40	64.5	22.1	499	10	Q9FWJ3	Q9fwj3 lycopersico
41	64.5	22.1	624	11	Q9D4B2	Q9d4b2 mus musculu
42	64.5	22.1	967	4	Q94336	Q94336 homo sapien
43	64.5	22.1	1038	5	O60983	O60983 dictyosteli
44	64	21.9	707	10	O04654	O04654 arabidopsis
45	64	21.9	770	5	Q9GPD7	Q9gpd7 drosophila

ALIGNMENTS

RESULT 1
Q9UDW8 ID Q9UDW8 PRELIMINARY; PRT; 615 AA.
AC Q9UDW8;
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DE WUGSC:H.DJ0747G18.3 PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=99063792; PubMed=9847074;
RA Sulston J.F., Waterston R.;
RT "Toward a complete human genome sequence."
RL Genome Res. 8:1097-1108(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Cordes M., Doela D.;
RT "The sequence of Homo sapiens PAC clone RP4-747G18.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC004876; AAD45830.1; -. 198097C5622AC087 CRC64;
SQ SEQUENCE 615 AA; 67257 MW; 198097C5622AC087 CRC64;

Query Match 85.4%; Score 249.5; DB 4; Length 615;
Best Local Similarity 88.5%; Pred. No. 5.1e-20;
Matches 54; Conservative 2; Mismatches 4; Indels 1; Gaps 1;
QY 1 QAEATROAAQEEERLADLASDLLLQYLQGGARQRLGGGRLQETQOERENR-EEAEQ 59
|||||
Db 310 QAEATROAAQEEERLADLASDLLLQYLQGGARQRLGGGRLQETQOERENR-EEAEQ 369

RP SEQUENCE FROM N.A.
 RC STRAIN=MAASTRICHT;
 RX MEDLINE=20366325; PubMed=10906222;
 RA Vink C., Beuken E., Bruggeman C.A.;
 RT "Complete DNA sequence of the rat cytomegalovirus genome.";
 RL J. Virol. 74:7656-7665(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAASTRICHT;
 RX MEDLINE=20473137; PubMed=11018281;
 RA Gruijthuisen Y.K., Beuken E., Bruggeman C.A., Vink C.;
 RT "Rat cytomegalovirus R89 is a highly conserved gene which expresses a
 RT spliced transcript.";
 RL Virus Res. 69:119-130(2000).
 DR EMBL; AF232689; AAF99114.1;
 SQ SEQUENCE 287 AA; 30152 MW; 353BE990EC8AF21B CRC64;

Query Match 25.7%; Score 75; DB 12; Length 287;
 Best Local Similarity 42.2%; Pred. No. 0.85;
 Matches 19; Conservative 7; Mismatches 11; Indels 8; Gaps 2;

QY 17 DLASDLLLQYLLQGGARQDLG-GRGLQETQOQERENEEAEQ 60
 Db 164 DVAGDV-----GGDRDDDCASGKKEAEAEAEAEAE 201

RESULT 5
 Q9CXK9 PRELIMINARY; PRT; 320 AA.
 AC Q9CXK9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DE 3200001K10RIK PROTEIN.
 GN 3200001K10RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Iaito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Adota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK014289; BAB29246.1;
 DR MGD; MGI:1919670; 3200001K10RIK.
 SQ SEQUENCE 320 AA; 36535 MW; 5DFD3BD47AD9EC5E CRC64;

Query Match 25.7%; Score 75; DB 11; Length 320;
 Best Local Similarity 38.0%; Pred. No. 0.95;
 Matches 19; Conservative 8; Mismatches 23; Indels 0; Gaps 0;

QY 6 RQAAQERLADLASDLLLQYLLQGGARQDLGGRGLQETQOQERENEE 55
 Db 31 RRAADEWDSELDLGEDLLSGKNQSDLSDELDNDLLQSDNEEE 80

RESULT 6
 Q98985 PRELIMINARY; PRT; 622 AA.
 AC Q98985;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CALNEXIN.
 OS Rana rugosa (Frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Ranidae; Rana.
 OX NCBI_TaxID=8410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96234004; PubMed=8654561;
 RA Yamamoto S., Makamura M.;
 RT "Calnexin: its molecular cloning and expression in the liver of the
 RT frog, Rana rugosa.";
 RL FEBS Lett. 387:27-32(1996).
 DR EMBL; D78590; BAAL1426.1;
 DR InterPro; IPR001580; Calreticulin.
 DR Pfam; PF00262; calreticulin; 1.
 DR PRINTS; PR00626; CALRETICULIN.
 DR PRODOM; PD001866; Calreticulin; 1.
 DR PROSITE; PS00804; CALRETICULIN_2; 1.
 DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
 SQ SEQUENCE 622 AA; 70802 MW; DAB3C807F1DEEE40 CRC64;

Query Match 25.2%; Score 73.5; DB 13; Length 622;
 Best Local Similarity 30.0%; Pred. No. 2.8;
 Matches 18; Conservative 10; Mismatches 25; Indels 7; Gaps 1;

QY 1 QAEATPQAAQEEERLADLASDLLLQYLLQGGARQDLGGRGLQETQOQERENEEAEQ 60
 Db 546 EEKAKKEADQEDNAEQAEK-----QTGEEGCAAGGSGQEEEEEEEEEE 598

RESULT 7
 O13098 PRELIMINARY; PRT; 897 AA.
 AC O13098;
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE MIDDLE MOLECULAR WEIGHT NEUROFILAMENT PROTEIN NF-M.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Gervasi C., Szaro B.G.;
 RL Brain Res. Mol. Brain Res. 0:0-0(0).
 CC -1- SIMILARITY: TO THE INTERMEDIATE FILAMENT FAMILY.
 DR EMBL; U85969; AAB53389.1;
 DR InterPro; IPR001664; IF.
 DR Pfam; PF00038; filament; 1.
 DR PROSITE; PS00226; IF; 1.
 KW Coiled coil; Intermediate filament.
 SQ SEQUENCE 897 AA; 100685 MW; D96221D9606DB077 CRC64;

Query Match 24.7%; Score 72; DB 13; Length 897;
 Best Local Similarity 31.1%; Pred. No. 6;
 Matches 19; Conservative 15; Mismatches 17; Indels 10; Gaps 3;

```

1 QAEATQAAQAEERLADSLDLLLYLQGGARQORDLGGRGLOETQOER-ENREFEAEQ 59
2 : || :: ||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::||| ::|||
3 585 EEEAEKEGEGEEKEEVAEE-----EGGEGEGEGEG--ETEEDKGEVEKEEAE 635
4
5 60 E 60
6 636 E 636
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1
```

DR PROSITE; PS00037; MYB_1; UNKNOWN_1
SQ SSQUENCE 540 AA; 61107 MW; 7FD6BB2E711COAFF CRC64;

Query Match 24.3%; Score 71; DB 11; Length 540;
Best Local Similarity 32.5%; Pred. No. 4.6;
Matches 25; Conservative 7; Mismatches 27; Indels 18; Gaps

OY 2 AENTROAAQ--BERLADIASLLLYLLOGGARQRDLGGRG-----IQ 43
| | | | | : | | | | | : | | | |
Db 156 ATATEHAQFQWPERLNNNVEELLQSLSIGKRDQQSSRRPQGQHKEQIOEHKLEEAQ 215
| | | | | : | | | | | : | | | |

OY 44 ETQOERENEREAEQE 60
| : : : | | : | | : | |
Db 216 EQBEQEEREEEEEAKOE 232

RESULT 10
Q9VK98 PRELIMINARY; PRT; 238 AA.

ID Q9VK98
AC Q9VK98;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DI 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)
DE ARET PROTEIN.
OS ARET OR CG6319.
GN Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.R., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berhan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:
RL "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).

```

DR EMBL; AE003635; AAF53180.2; -.
DR HSP; P09651; IUP1.
DR FlyBase; FBgn0000114; aret.
SQ SEQUENCE 238 AA; 28069 MW; FBBA07BABC2D1805 CRC64;

Query Match          24.0%; Score 70; DB 5; Length 238;
Best Local Similarity 34.9%; Pred. No. 2.5;
Matches 22; Conservative 13; Mismatches 12; Indels 16; Gaps 3;

QY 6 RQAAAEERLADLAS-----DLLLQYLQGGARQDLGGRL---QETQQRERERE 57
DB 149 QQTSTYQQQLSQAALQDHDLL-----QTQELFHKQLATMQEYQRERERER 200
QY 58 EQE 60
DB 201 ERE 203

P T 11
O 3
ID O96133 PRELIMINARY; PRT; 1979 AA.
AC O96133;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 237.7 KDA PROTEIN.
GN PFB0145C.
OS Plasmodium falciparum.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=5833;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95021743; PubMed=9804551;
RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L.,
RA Koonin E.V., Shallom S., Mason T., Yu K., Fujii C., Pederson J.,
RA Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M.,
RA Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O.,
RA Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;
RT "Chromosome 2 sequence of the human malaria parasite Plasmodium
RT falciparum";
RL Science 282:1126-1132(1998).
DR EMBL; AE001375; AAC71819.1; -.
KW Hypothetical protein.
SQ SEQUENCE 1979 AA; 237745 MW; 5C6CD6307AFF7D37 CRC64;

Query Match          24.0%; Score 70; DB 5; Length 1979;
Best Local Similarity 44.7%; Pred. No. 23;
Matches 17; Conservative 9; Mismatches 10; Indels 2; Gaps 1;

QY 23 LLYLQGGARQDLGGRLQETQQREREREAEQE 60
DB 314 LLEYEEKNQIDLNLK--QEKEKEKEKEKEKE 349

RESULT 12
Q9L2C3 PRELIMINARY; PRT; 1326 AA.
ID Q9L2C3
AC Q9L2C3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE LARGE ALA/GLU-RICH PROTEIN.
GN SC8f4.01c.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;

```

```

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL; AL137242; CAB70627.1;
SQ SEQUENCE 1326 AA; 146433 MW; 9C0DCB984863FA38 CRC64;

Query Match          23.8%; Score 69.5; DB 2; Length 1326;
Best Local Similarity 37.7%; Pred. No. 17;
Matches 23; Conservative 7; Mismatches 26; Indels 5; Gaps 2;

QY 1 QAEATRQAAQEERLADLASDLLLQYLQGGARQR---DLGGRLQETQQRERERE 57
DB 682 RVRAEAQTAA--ERIAAEASEALAAQEEAARRRRREABELLGSAQEADQERERV 739
QY 58 E 58
DB 740 E 740

RESULT 13
O18409 PRELIMINARY; PRT; 808 AA.
ID O18409
AC O18409;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17; Last annotation update)
DE TESTIS-SPECIFIC RNP-TYPE RNA BINDING PROTEIN.
GN ARET OR BRUNO OR CG6319.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Webster P.J., Liang L., Lasko P., MacDonald P.M.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U73846; AAB61993.1; -.
DR HSP; P19339; 2SXL.
DR FlyBase; FBgn0000114; aret.
DR InterPro; IPR000504; RRM.
DR Pfam; PF00076; Rrm; 3.
DR SMART; SM00360; RRM; 3.
DR PROSITE; PS0102; RRM; 3.
DR PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
SQ SEQUENCE 808 AA; 88088 MW; 55DAD7C50F9BADC6 CRC64;

Query Match          23.6%; Score 69; DB 5; Length 808;
Best Local Similarity 34.9%; Pred. No. 12;
Matches 22; Conservative 12; Mismatches 13; Indels 16; Gaps 3;

QY 6 RQAAAEERLADLAS-----DLLLQYLQGGARQDLGGRL---QETQQRERERE 57
DB 149 QQTSTYQQQLSQAALQDHDLL-----QTQELFHKQLATMQEYQRERERER 200
QY 58 EQE 60
DB 201 ERE 203

```

```
RESULT 14
D P79805 PRELIMINARY; PRT; 773 AA.
C P79805;
T 01-MAY-1997 (TReMBLrel. 03, Created)
T 01-MAY-1997 (TReMBLrel. 03, Last sequence update)
T 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
E MICROTUBULE-BASED MOTOR PROTEIN.
N FKIF2.
S Morone saxatilis (Striped bass).
C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
C Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
C Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
C Moronidae; Morone.
X NCBI_TaxID=34816;
N [1]
P SEQUENCE FROM N.A.
C TISSUE=RETINA;
A Bost-Usinger L.;
I submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
X MBL; U64819; AAB39558.1; -.
R HSSP; P17119; 3KAR.
R InterPro: IPR001752; kinesin.
R Pfam: PF00225; kinesin; 1.
R PRINTS; PR00380; KINESINHEAVY.
R SMART; SM00129; KISC; 1.
R PROSITE; PS00411; KINESIN_MOTOR_DOMAIN1; 1.
R PROSITE; PS00067; KINESIN_MOTOR_DOMAIN2; 1.
W ATP-binding; Coiled coil; Microtubules; Motor protein.
Q SEQUENCE 773 AA; 87392 MW; B16832F0A4B0FA45 CRC64;

Query Match 23.5%; Score 68.5; DB 13; Length 773;
Best Local Similarity 35.5%; Pred. No. 13;
Matches 22; Conservative 10; Mismatches 17; Indels 13; Gaps 2;

y 7 QAAQAQERLADLASDL-----LQYLQGGARQDILGGRGLQETQQREREREAE 58
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
b 73 QARQRRRLAELEKSSRAESRLRQDLESTQIDDSRLQLEEAR-----RQKEAE 127
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
y 59 QE 60
|||
b 128 OK 129

RESULT 15
P9U412 PRELIMINARY; PRT; 3469 AA.
A J9U412;
T 01-MAY-2000 (TReMBLrel. 13, Created)
T 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
T 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
E SANT DOMAIN PROTEIN SMRTER.
N SMR OR SMRTER OR CG4013.
S Drosophila melanogaster (Fruit fly).
C Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
C Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
C Ephydroidea; Drosophilidae; Drosophila.
X NCBI_TaxID=7227;
N [1]
P SEQUENCE FROM N.A.
X MEDLINE=99417957; PubMed=10488333;
A Tsai C.-C., Kao H.-Y., Yao T.-P., McKeown M., Evans R.M.;
T "SMRTER, a Drosophila nuclear receptor coregulator, reveals that Ecr-
mediated repression is critical for development.";
L Mol. Cell 4:175-186(1999).
R EMBL; AF175223; AADS2614.1; -.
R FlyBase; FBgn0024308; Smr.
R InterPro; IPR002086; Aldehyde_dehydr.
R InterPro; IPR001005; Myb_DNA_bind.
R Pfam; PF00249; myb_DNA-binding; 1.
R SMART; SM00395; SANT; 1.
```

DR PROSITE: PS00687; ALDEHYDE_DEHYDR_GLU: UNKNOWN_1.
SQ SEQUENCE 3469 AA; 364115 MW; 6284E14C5C247CD9 CRC64;

Query Match 23.5%; Score 68.5; DB 5; Length 3469;
Best Local Similarity 31.7%; Pred. No. 60;
Matches 19; Conservative 12; Mismatches 26; Indels 3; Gaps 1;
QY 1 QAEATROAAQAQEEERLADLASDLLLQYLQGGARQDILGGRGLQETQQREREREAEQE 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2707 QVQQQQQQQQQQQHNALEDRDREMRERERE---REREREQRRERERERE 2763

Search completed: January 15, 2002, 17:33:27
Job time: 440 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:20:47 ; Search time 12.6 Seconds
(without alignments)
107.159 Million cell updates/sec

Title: US-09-619-198-7

Perfect score: 292

Sequence: 1 QAEATRQAAQEEERLADLAS.....GLQETQOQRENEREEAEQE 60

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

1 number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/6C_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66.5	22.8	651	3	US-08-650-766-6
2	66.5	22.8	651	3	US-08-922-635-5
3	66.5	22.8	1070	3	US-08-922-635-22
4	62.5	21.4	515	2	US-08-705-660-46
5	62.5	21.4	515	3	US-08-989-045-46
6	62.5	21.4	1185	4	US-09-041-886-23
7	61.5	21.1	414	5	PCT-US93-03077-3
8	61.5	21.1	1093	5	PCT-US93-03077-1
9	61	20.9	903	2	US-08-853-310-2
10	59	20.2	240	2	US-08-114-555A-8
11	59	20.2	240	3	US-08-559-397A-14
12	58	19.9	1162	2	US-08-728-323A-2
13	57.5	19.7	450	2	US-08-665-037-2
14	57.5	19.7	450	2	US-08-666-067-2
15	57.5	19.7	450	2	US-08-732-870-2
16	57	19.5	1958	1	US-07-945-283-2
17	56.5	19.3	76	1	US-08-253-155A-28
18	56.5	19.3	77	1	US-08-253-155A-47
19	56.5	19.3	83	1	US-08-253-155A-38
20	56.5	19.3	240	2	US-08-114-555A-6
21	56.5	19.3	240	3	US-08-559-397A-12
22	56.5	19.3	521	4	US-09-370-368-4
23	56.5	19.3	3111	2	US-08-460-309-4
24	56.5	19.3	3111	2	US-08-125-077-4
25	56	19.2	234	4	US-09-040-483-5
26	56	19.2	905	2	US-08-574-959A-9
27	56	19.2	905	4	US-09-357-014-9

28	56	19.2	955	1	US-08-006-676B-1	Sequence 1, Appli
29	56	19.2	955	1	US-08-282-845-2	Sequence 2, Appli
30	56	19.2	955	2	US-08-428-414A-3	Sequence 3, Appli
31	56	19.2	955	5	PCT-US94-00324-1	Sequence 1, Appli
32	56	19.2	1135	2	US-08-574-959A-7	Sequence 7, Appli
33	56	19.2	1135	4	US-09-357-014-7	Sequence 7, Appli
34	55.5	19.0	316	4	US-08-098-327E-31	Sequence 31, Appli
35	55	18.8	162	1	US-08-244-557-2	Sequence 2, Appli
36	55	18.8	304	4	US-09-042-225-7	Sequence 7, Appli
37	54.5	18.7	55	1	US-08-199-776-4	Sequence 4, Appli
38	54.5	18.7	55	3	US-08-663-731-4	Sequence 4, Appli
39	54.5	18.7	55	3	US-08-879-338-4	Sequence 4, Appli
40	54.5	18.7	55	5	PCT-US95-02044-4	Sequence 4, Appli
41	54.5	18.7	1178	1	US-08-199-776-2	Sequence 2, Appli
42	54.5	18.7	1178	3	US-08-663-731-2	Sequence 2, Appli
43	54.5	18.7	1178	3	US-08-879-338-2	Sequence 2, Appli
44	54.5	18.7	1178	5	PCT-US95-02044-2	Sequence 2, Appli
45	54	18.5	161	5	PCT-US95-05741-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-08-650-766-6
; Sequence 6, Application US/08650766D
; Patent No. 6015690
; GENERAL INFORMATION:
; APPLICANT: PILETZ, John E.
; APPLICANT: IVANOV, Tina R.
; TITLE OF INVENTION: DNA SEQUENCE ENCODING A HUMAN IMIDAZOLINE RECEPTOR AND
; TITLE OF INVENTION: METHOD FOR CLONING THE SAME
; FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6015690
; CURRENT APPLICATION NUMBER: US/08/650.766D
; CURRENT FILING DATE: 1996-05-20
; EARLIER APPLICATION NUMBER: US 60/012.600
; EARLIER FILING DATE: 1996-03-01
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-650-766-6

Query Match 22.8%; Score 66.5; DB 3; Length 651;
Best Local Similarity 30.6%; Pred. No. 1.6;
Matches 19; Conservative 12; Mismatches 22; Indels 9; Gaps 3;
QY 1 QAEATRQAAQEEERLADLASLLLOYLQ--GARQDLGGRGLOETQOQRENEREEAE 58
Db 220 QGEGEEDEEEEDVAEN---RYFEMGPPDVEERGGQG-----EEEEEEDEEAE 272
QY 59 QE 60
Db 273 EE 274

RESULT 2

US-08-922-635-5
; Sequence 5, Application US/08922635A
; Patent No. 6033871
; GENERAL INFORMATION:
; APPLICANT: PILETZ, John E.
; APPLICANT: IVANOV, Tina R.
; TITLE OF INVENTION: DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
; TITLE OF INVENTION: AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: Corrected Sequence Listing
; Patent No. 6033871
; CURRENT APPLICATION NUMBER: US/08/922.635A
; CURRENT FILING DATE: 1997-09-03

ATTORNEY/AGENT INFORMATION:

```

EARLIER APPLICATION NUMBER: 08/650,766
EARLIER FILING DATE: 1996-05-20
EARLIER APPLICATION NUMBER: 60/012,600
EARLIER FILING DATE: 1996-03-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 651
TYPE: PRT
ORGANISM: Homo sapiens
1-08-922-635-5

Query Match      22.8%; Score 66.5; DB 3; Length 651;
Best Local Similarity 30.6%; Pred. No. 1.6;
Matches 19; Conservative 12; Mismatches 22; Indels 9; Gaps 3;

      1 QEAETQAAQAEERLADSLILQYLQGG--GARQDLDGGRGLQETQOERENREERAE 58
      220 QGQGEDEDEEEEDVAEN---RYFEMGPPDVEEEGGGQG---EEEEEEDEDEAE 272

      59 QE 60
      273 EE 274

RESULT 3
3-08-922-635-22
Sequence 22, Application US/08922635A
Patent No. 6033871
GENERAL INFORMATION:
APPLICANT: PILETZ, John E.
APPLICANT: IVANOV, Tina R.
TITLE OF INVENTION: DNA MOLECULES ENCODING IMIDALINE RECEPTIVE POLYPEPTIDES
TITLE OF INVENTION: AND POLYPEPTIDES ENCODED THEREBY
FILE REFERENCE: Corrected Sequence Listing
Patent No. 6033871
CURRENT APPLICATION NUMBER: US/08/922,635A
CURRENT FILING DATE: 1997-09-03
EARLIER APPLICATION NUMBER: 08/650,766
EARLIER FILING DATE: 1996-05-20
EARLIER APPLICATION NUMBER: 60/012,600
EARLIER FILING DATE: 1996-03-01
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 1070
TYPE: PRT
ORGANISM: Homo sapiens
--922-635-22

Query Match      22.8%; Score 66.5; DB 3; Length 1070;
Best Local Similarity 30.6%; Pred. No. 3;
Matches 19; Conservative 12; Mismatches 22; Indels 9; Gaps 3

      1 QEAETQAAQAEERLADSLILQYLQGG--GARQDLDGGRGLQETQOERENREERAE 58
      198 QGQGEDEDEEEEDVAEN---RYFEMGPPDVEEEGGGQG---EEEEEEDEDEAE 250

      59 QE 60
      251 EE 252

RESULT 4
S-08-705-660-46
Sequence 46, Application US/08705660
Patent No. 5858683
GENERAL INFORMATION:
APPLICANT: KEESEE, SUSAN
APPLICANT: OBAR, ROBERT
APPLICANT: WU, YING-JYE

```